More than the sum of its parts: Coarse-grained peptide-lipid interactions from a simple cross-parametrization

Tristan Bereau,[a] Zun-Jing Wang,[b] and Markus Deserno[c]
Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA

(Received 8 November 2013; accepted 4 February 2014; published online 17 March 2014)

Interfacial systems are at the core of fascinating phenomena in many disciplines, such as biochemistry, soft-matter physics, and food science. However, the parametrization of accurate, reliable, and consistent coarse-grained (CG) models for systems at interfaces remains a challenging endeavor. In the present work, we explore to what extent two independently developed solvent-free CG models of peptides and lipids—of different mapping schemes, parametrization methods, target functions, and validation criteria—can be combined by only tuning the cross-interactions. Our results show that the cross-parametrization can reproduce a number of structural properties of membrane peptides (for example, tilt and hydrophobic mismatch), in agreement with existing peptide-lipid CG force fields. We find encouraging results for two challenging biophysical problems: (i) membrane pore formation mediated by the cooperative action of several antimicrobial peptides, and (ii) the insertion and folding of the helix-forming peptide WALP23 in the membrane. © *2014 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4867465]

I. INTRODUCTION

A number of fascinating phenomena occur at the interface between two or more phases.[1–3] The local inhomogeneities that establish the structure of the interface, which underly the emergent characteristics of that interface relevant at much larger dimensions, require the consideration and/or bridging of a broad range of scales: the accurate description of each individual phase sets a lower bound on the overall system size, while the equilibration *between* phases takes time. Though all-atom molecular dynamics simulations have recently shown capable to reach system sizes in the 10^6 atom range[4–7] and remarkably long time scales for smaller systems,[8] the combination of both large length- and time-scales remains a problematic issue in view of adequate sampling. In this regard, coarse-grained (CG) models, which average over certain degrees of freedom to provide a simplified representation of the system, are enjoying ever-growing attention in (bio)molecular fields.[9–12]

Irrespective of the resolution, many force-field parametrization methodologies aim at reproducing target properties of an underlying microscopic model. These might be thermodynamic data from bulk systems (such as pure-liquid density and heat of vaporization for atomistic models[13]), or local interaction statistics (such as structure factors or pair forces[14]). While they can help reproduce a number of thermodynamic properties, the model's transferability across both thermodynamic parameters (e.g., temperature, pressure) and environments (e.g., bulk vs. interface) often remains of concern. In this regard, the development of CG models suitable for interfacial systems is a challenging effort. Recently, Jochum *et al.*[15] tackled the CG parametrization of inhomogeneous systems: they generalized structure-based coarse-graining methods for slabs of liquids, which both improved thermodynamic properties at the interface and preserved the structure in the bulk.

Among interfacial systems, peptide–lipid interactions have generated increasing interest in the biochemistry and biophysics communities. Several lipid-protein CG models of various degrees of resolution have been developed and studied in the past.[16–25] Though different in resolution, accuracy, and transferability, they reproduce—possibly by construction of the model—a number of key features that make protein–lipid interactions remarkable, such as hydrophobic mismatch, protein insertion, and protein-driven pore formation. For most of these force fields, peptide and lipid molecules were parametrized consistently by deriving them either simultaneously or against one another (for example, by using the same overall CG scheme, as is done in the MARTINI model[21,26]).

In the present work, we aim at studying the benefit one achieves by linking two *existing* and independently derived models of peptides and lipids. Since the two models—described below—were designed with different mapping schemes, parametrization methods, target functions, and validation criteria, what phenomenological aspects of peptide–lipid interactions can be rescued by merely tuning the cross-interactions? In this sense, the final model does not aspire to providing complete transferability (e.g., across protein sequences, lipid composition), but rather qualitative aspects of peptide–lipid interactions from a simple cross-parametrization. Given how technically involved a fully self-consistent parametrization of a multi-component system is, a question of considerable practical relevance is: How much can one gain from the easier (even though still work-intensive) approach of post-parametrizing cross-terms? The question is

[a]Present address: Max-Planck-Institut für Polymerforschung, Ackermannweg 10, 55128 Mainz, Germany. Electronic mail: bereau@mpip-mainz.mpg.de
[b]Electronic mail: zunjingwang@gmail.com
[c]Electronic mail: deserno@andrew.cmu.edu

certainly not trivial: Mullinax and Noid[27] have shown that self-consistent parametrizations of methanol (M) and neopentane (N) yield different M–M and N–N interaction potentials across different mixing ratios. In other words, compared to a simple cross-parametrization, an optimal force field will likely have different CG interactions between the subcomponents among themselves. Other examples have shown more promising: a recent study from DeMille *et al.*[28] presented a cross-parametrization of an implicit CG DNA model and a water/ions model. This issue has proven much more pronounced for CG than atomistic models—the main reason being that CG potentials in fact contain information about the environment and degrees of freedom that are coarse-grained out, and as such are *effective* pair potentials.

In this study, we will approach this question with a specific example from the field of protein–lipid biophysics: We will parametrize the protein–lipid cross-interactions between two existing solvent-free CG models of roughly comparable resolution by reproducing the energetics of the insertion of single amino acids into a 1, 2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) bilayer, as described in Sec. III. Subsequently, this cross-parametrization will be tested on simple structural properties of transmembrane proteins, such as tilt angle, helix–helix distance, and their crossing angle. We then apply the model to two challenging situations: First, we follow the creation of a trans-bilayer pore by eight antimicrobial peptides (magainin) initially outside the membrane; second, we follow the insertion and folding of a transmembrane helical peptide (WALP).

II. INDIVIDUAL MODELS

Let us briefly describe the separate peptide- and lipid-models we base this study on. The implicit-solvent CG protein model used here[29] accounts for amino-acid specificity and is capable of representing secondary structure without explicitly biasing the force field toward any particular conformation (native or not). It represents an amino acid by four beads, three of which are devoted to the backbone and the last one representing the side chain (Fig. 1(a)). The model was built using

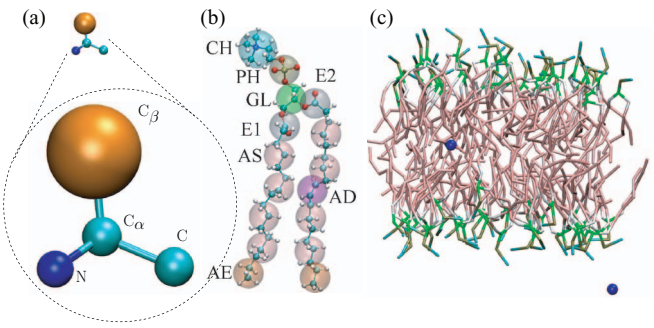

FIG. 1. (a) CG amino acid (colors: N dark blue, C_α and C′ cyan, C_β orange) to scale with the lipid (top) and magnified (bottom). (b) CG POPC lipid.[35] Reproduced in part with permission from Wang and Deserno, J. Phys. Chem. B **114**(34), 11207–11220 (2010). Copyright 2010 American Chemical Society. (c) Side view of a 72-molecule DOPC bilayer as well as two amino acid side chains (in blue) constrained at the center and outside of the membrane. Visualizations rendered with VMD.[46]

a top-down scheme in which phenomenological interactions are incorporated to reproduce different properties of the system (e.g., hydrophobicity or hydrogen bonds). These interactions were parametrized to (i) qualitatively reproduce the Ramachandran plot of simple tripeptides and (ii) systematically fold a *de novo* three-helix bundle. After parametrization, the model was shown capable of folding several simple helical proteins and reproduce an oligopeptide aggregation scenario using the same force field parameters and without any primary sequence dependent (custom-made) bias.[29] While this model handles a variety of secondary structure formation aspects very accurately, for instance, cooperativity in helix and helix-bundle folding transitions[30,31] and β-barrel formation,[32] it is not detailed enough to achieve folding of complex protein structures. However, nontrivial tertiary or quaternary folds can still be treated by adding stabilizing elastic networks,[33,34] but unstructured tails and flexible loops can now be left free to explore a meaningful phase space.

The implicit-solvent CG lipid model used here[35,36] provides a similar resolution as the MARTINI force field (about three to four heavy atoms per bead). Fig. 1(b) represents a coarse-grained 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) lipid.[35] It is constructed from 16 beads and 8 different bead types: CH Choline; PH Phosphate; GL Glycerol; E1 and E2 Ester groups; AS Saturated alkyl group $-(CH_2–CH_2–CH_2)–$; AD Unsaturated alkyl group $-(CH_2{=}CH_2)–$; and AE Hydrocarbon endgroup $-(CH_2–CH_3)$. It was systematically parametrized to reproduce the radial distribution functions between the corresponding chemical moieties as measured within an atomistic simulation of POPC using iterative Boltzmann inversion.[37,38] The implicit nature of the solvent in the CG system does not automatically provide a hydrophobic effect, which is compensated by an additional effective cohesion.[35] Besides its ability to self-assemble a random dispersion of lipids into a bilayer, the model can almost quantitatively reproduce many of the properties of a POPC bilayer, including the bending and stretching moduli, mass density profile, and orientational P_2 order parameters of intramolecular bonds. The construction of *other* types of lipids (e.g., DOPC, DPPC—1, 2-dipalmitoyl-*sn*-glycero-3-phosphocholine)—starting from the same set of coarse-grained beads as POPC—has shown transferability in terms of structure, area per lipid, and the temperature dependence of the main phase transition.[36]

III. CROSS-PARAMETRIZATION

In this section, we derive potentials for the interactions between the protein and the lipid force fields. Our cross-parametrization aims at reproducing the potential of mean force (PMF) for the insertion of individual amino acid side chains into a DOPC bilayer. These PMFs provide additional spatial resolution compared to experimental hydrophobicity scales.[39,40] Because of the planar structure of lipid bilayers, free energies as a function of the z-coordinate (i.e., perpendicular to the bilayer plane), $F(z)$, incorporate much of the thermodynamic information required to describe the energetics of insertion of molecules in a membrane. Moreover, the spatial ordering of lipid groups (such as alkyl chains or the

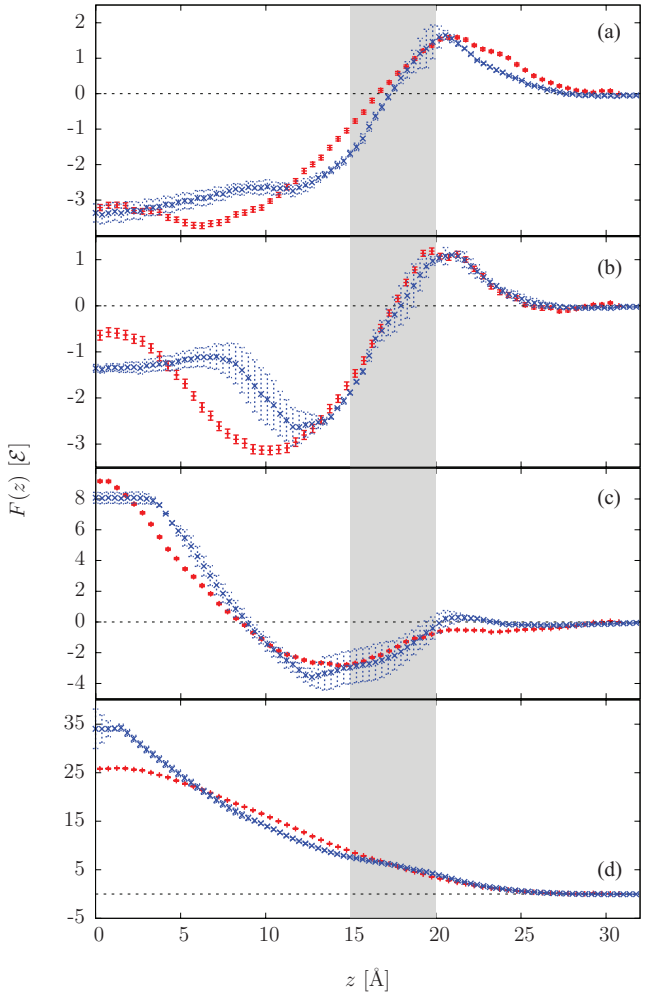

FIG. 2. Free-energy profiles for the insertion of a single amino acid in a DOPC bilayer: CG (red, solid) and atomistic (blue, dashed; from MacCallum *et al.*[41]) (a) Ala; (b) Cys; (c) Gln; and (d) Glu^-. Shaded regions roughly indicate the interfacial region (i.e., glycerol backbone and phosphate); smaller z values correspond to hydrophobic tails while larger z values denote the lipid head region and aqueous environment. PMFs for all other amino acids are reported in Figs. S1– S3 of the supplementary material.[44]

phosphate moiety) in a bilayer provides a means to understanding the impact of each group onto $F(z)$.

The reference data are based on atomistic simulations performed by MacCallum *et al.*:[41] amino acid side chains (starting at the β-carbon) were inserted in a 64-molecule DOPC bilayer using the OPLS all-atom force field.[42,43] Their study provided curves for the free energy of insertion of side chains for all standard amino acids except Gly, His, and Pro. While Gly and Pro were not calculated because of the chemistry of their side chain (the side chain of Gly holds no heavy atom; the side chain of Pro is connected to the backbone at *both* the central carbon and the amide group), His causes issues due to its multiple protonation states. Ionizable residues Arg, Lys, Asp, and Glu were calculated for both the charged and neutral forms. The results are reproduced in Fig. 2 for a selected list of residues, and in Figs. S1– S3 of the supplementary material[44] for all the others.

The cross-parametrization thus consists of optimizing potentials of interaction between lipids and amino acids to reproduce the PMFs derived from MacCallum *et al.*[41] Unfortunately, these free-energy profiles only provide partial information on the system. For instance, the PMFs derived from atomistic simulations only contain information on the interaction between lipids and side chains, rather than the entire amino acid. The atomistic PMFs will be used as target function for the free energy of insertion of the CG side-chain bead C_β in a DOPC bilayer. Purely repulsive interactions between lipid and peptide C_α, N, and C′ beads will model the excluded volume effect between lipids and the protein backbone. See Fig. 1(a) for details. Common to many force-field parametrization endeavors, we point out that the present parametrization process is an underdetermined problem, such that there is no unique solution. In particular, a single target function $F(z)$ cannot encode the information for multiple interaction potentials between different lipid moieties and the amino acid, and hence one must supplement the cross-parametrization with additional physical constraints. This is in contrast to more conventional instances of structure based coarse-graining, where each interaction on the CG level has one atomistically derived pair correlation function to match, and then the Henderson theorem[45] guarantees (at least in principle) a unique solution.

A. Simulation and analysis methods

Each simulation consisted of a 72-molecule DOPC bilayer as well as *two* amino acids, one inserted in each leaflet (as in MacCallum *et al.*[41]). This has two advantages: First, it doubles the statistics on $F(z)$ achievable in each simulation. Second, and maybe more importantly, inserting an amino acid in only one of the two leaflets will compress this leaflet and stretch the opposing one. This creates an elastic stress which cannot relax even if one ensures that the overall mechanical tension vanishes. It is easy to see that this would lead to a finite size correction of the free energy which scales inversely with bilayer area. See Fig. 1(c) for a snapshot of the simulation setup.

The CG units used throughout, as well as most simulation parameters and protocols (except when explicitly mentioned), follow Bereau and Deserno[29] and Wang and Deserno.[35] The difference between the two models' time units, τ, can complicate the interpretation of dynamical processes, but since this goes beyond the scope of the present article, we will not discuss the associated issues. Simulations were run at constant temperature ($T = T_{\text{body}} = 310\,\text{K}$) and zero lateral bilayer tension ($\Sigma = 0$) for a total simulation time $t = 200\,000\,\tau$. A time step $\delta t = 0.1\,\tau$ was used for amino-acid side-chain insertion simulations.

The distance between the amino-acid side-chain and the bilayer midplane was calculated by measuring the difference between (i) the z-coordinate of the side chain bead and (ii) the z-coordinate of the center of mass of the bilayer, which we will identify with the bilayer midplane. A total of 10 000 data points were recorded at intervals separated by $20\,\tau$. Umbrella sampling[47] was used on the z-positions of the two inclusions, applying a harmonic restraint of spring constant $k = 2\,\mathcal{E}/\text{Å}^2$, where $\mathcal{E} = k_B T_{\text{body}} \approx 0.6$ kcal/mol is the unit of energy of the two models. Pairs of two inclusions were systematically

placed 30 Å apart (vertically; corresponding roughly to the height of a monolayer) to avoid artifacts.[48] A total of 32 such umbrellas, each separated by 1 Å, yielded biased distributions in an interval $0 < z < 32$ Å, where $z = 0$ corresponds to the bilayer midplane. We unbiased the sampled distributions using the Weighted Histogram Analysis Method (WHAM)[49–51] and calculated error bars from bootstrapping.[52]

B. Interaction potentials and parametrization

Potentials of interaction were set between all lipid and amino acid bead types using only Lennard-Jones (LJ) and Weeks-Chandler-Andersen (WCA) potentials as functional forms

$$U_{\mathrm{LJ}}(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right],$$

$$U_{\mathrm{WCA}}(r) = \begin{cases} 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} + \frac{1}{4}\right], & \text{if } r < 2^{1/6}\sigma \\ 0, & \text{otherwise.} \end{cases}$$

$U_{\mathrm{LJ}}(r)$ was consistently cut and shifted to 0 at a distance $r = 15$ Å. During parametrization, we systematically assigned a fixed interaction strength $\epsilon = 1.0\,\mathcal{E}$ for WCA potentials, since the function is largely insensitive to changes in this parameter.

Lipid bead types AS and AD were not distinguished in terms of their interactions with peptide beads. They were both denoted as AS in the following. Likewise, E1 and E2 were both associated to the same bead type ES in terms of nonbonded interactions with amino acids.

Sets of interaction potentials were iteratively refined until the resulting PMF matched the target curve. More details on the parametrization protocol can be found in the supplementary material.[44]

C. Optimal parameters and PMFs

1. Side chains

Table S1 in the supplementary material[44] shows the sets of interaction potentials and parameters that we arrived at using the iteration procedure outlined above. For each amino-acid–lipid-group pair, the table displays (i) the functional form of the interaction; (ii) the strength of the interaction, ϵ; and (iii) the excluded volume of the interaction, σ. Fig. 2 shows both the CG (red) and atomistic (blue) PMFs for a subset of amino acids: Ala, Cys, Gln, and Glu^-. All PMFs strongly resemble one of the four profiles shown in Fig. 2. The accuracy (i.e., the agreement between CG and atomistic) shown here is representative of what has been achieved for the entire set of PMFs—displayed in Figs. S1– S3 of the supplementary material[44]—and comparable with the accuracy with which MARTINI reproduces the same atomistic PMFs.[21]

We measured the deviation between CG and AA PMFs by computing a root-mean-squared error (RMSE) that was averaged over a large number of points along the PMF ($0 < z < 32$). The RMSE of each amino acid, shown in the supplementary material,[44] ranges between 0.3 and 3 $k_{\mathrm{B}}T$. It is highest for charged amino acids where the CG model fails to reproduce steep behavior close to the bilayer mid-plane (see below). The RMSE averaged over all amino acids is 1.0 $k_{\mathrm{B}}T$. We discuss in detail our results in the supplementary material.[44] The resulting parameters semi-quantitatively capture the important energetic features of amino side chains interacting with a lipid membrane, on par with the MARTINI force field.

2. Backbone

To reproduce the excluded volume effect between lipids and the protein backbone, purely repulsive interactions (WCA) were set between all lipid bead types and peptide backbone particles N, C_α, and C′ with parameters $\epsilon = \epsilon_{\mathrm{bb}} = 0.02\,\mathcal{E}$ (see Ref. 29 for details) and $\sigma = \sigma_{\mathrm{peptide}} + \sigma_{\mathrm{lipid}}$, where $\sigma_{\mathrm{peptide}}$ corresponds to the van der Waals radius of peptide backbone particles N, C_α, or C′,[29] and $\sigma_{\mathrm{lipid}} = 3.0$ Å represents an average lipid bead radius. While a more detailed parametrization would seem possible, the chosen procedure is sufficient to reproduce the main steric effects.

D. Parametrization of glycine, histidine, and proline

Due to the lack of atomistic data for residues Gly, His, and Pro, their PMF curves were extrapolated from other amino acids similar in chemical structure or hydrophobicity. Details of this construction—based on chemical or hydrophobic similarities to other amino acids—are provided in the supplementary material.[44] The resulting reference curves are shown in Figs. S1– S3, as well as the corresponding optimized CG curves. Coarse-grained parameters used to reproduce the target data are shown in Table S1.

E. N- and C-termini of proteins

Let us conclude with a few remarks on the peptide's N- and C-termini. While this issue could have been addressed previously in the presentation of the CG protein model,[29] their effects on the conformations sampled is often rather limited, and so the termini are usually neglected in coarse-grained peptide models parametrized in an aqueous environment. However, these termini play an important role for transmembrane proteins for which most amino acids are hydrophobic, but the termini are either polar or charged. This scenario ensures that the protein is *integral* to the membrane, i.e., it spans the bilayer thickness (rather than "dive" into the membrane if all amino acids were hydrophobic). Common examples of N- and C-termini include the acetyl and *n*-methyl amide groups, respectively (Figs. S5(e) and S5(f) of the supplementary material[44]). We devised a simple, phenomenological model of N- and C-termini by modeling them as Gly residues—providing flexibility in the backbone while preventing side chain-side chain interactions—with interactions that are repulsive enough to keep a transmembrane peptide integral to the membrane. More details can be found in the supplementary material.[44]

IV. STRUCTURAL PROPERTIES

In the following, we perform a set of simulations of inserted residues, transmembrane, and antimicrobial proteins

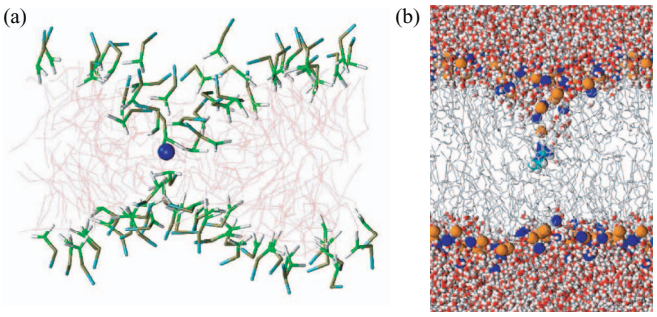

FIG. 3. (a) CG configuration of Arg^+ (blue sphere) in a DOPC bilayer (thin lines: hydrocarbon tails; licorice: head groups). The insertion of Arg^+ creates a strong, local deformation of the bilayer. (b) Atomistic configuration of Arg^+ in a DOPC bilayer. Reprinted with permission from MacCallum *et al.*, Biophys. J. **94**, 3393–3404 (2008), Copyright 2008 Biophysical Society.

interacting with bilayer membranes. The results are compared with available experimental and atomistic simulation data to assess the cross-parametrization's performance. The choice of test systems is to a large extent motivated by the validation checks used by existing CG models of similar spirit.[21,24]

A. Water defects in a solvent-free model

The partitioning of polar and charged residues into the hydrocarbon region of the bilayer have been associated with large water defects[41,53,54] (see Fig. 3(b)). MacCallum *et al.*[41] observed the stabilization of narrow pores that allow water molecules to interact with a polar/charged residue. Their simulations suggest that such a water channel persists even when a charged Arg residue is located in the bilayer midplane.[41]

Fig. 3 shows that in our CG cross-parametrization Arg^+ also strongly deforms the bilayer locally. The local ordering and packing of the lipids is strongly perturbed because the Arg^+ residue interacts more favorably with the lipid head groups than the hydrocarbon tails (see Fig. S1 in the supplementary material[44]). Overall, we observe a localized thinning of the bilayer—a remarkable effect considering the absence of explicit electrostatics. It is difficult to achieve the same type of water defect in MARTINI—not because of any principal difficulties of matching the PMF, but because the coarse-grained MARTINI water (with 4 waters subsumed into a single CG bead) tends to be too bulky to penetrate into this localized narrow defect. This was illustrated in a study from Vorobyov *et al.*,[55] where the free-energy of insertion of Arg^+ and the coordination number of water molecules was measured at both atomistic (polarizable and non-polarizable) and CG levels. Since our model is solvent-free, this particular complication does not exist.

B. Protein fluctuations in water and in the membrane

A peptide that folds into an α-helix both in water and in the membrane will likely *not* have the same flexibility in the two environments: the free energy of breaking a peptide–peptide hydrogen bond is higher in the membrane because there are no available hydrogen-bond donors/acceptors in the apolar solvent of lipid tails. One thus expects a helix to be stiffer in the membrane. This is indeed what was observed from atomistic simulations for alamethicin (Alm), a channel-forming, fungal peptide (Tieleman *et al.*;[56] sequence shown in Table I): root-mean-square fluctuations (RMSF)[59,60] of the helix in water and in a POPC bilayer (reproduced in Fig. 4; denoted "water AA" and "membrane AA," respectively) clearly make the point. It is not clear, though, whether the same should hold at the CG level:

- On the one hand, the strength of the hydrogen-bond interaction ϵ_{hb} was parametrized to reproduce the structure of helical proteins in *water*.[29] This suggests the need for a reparametrization of ϵ_{hb} in a membrane environment (the value would likely go *up* to reproduce the above-mentioned change in free energy[61]).
- On the other hand, our CG model couples an *implicit*-water solvent with an *explicit*-membrane environment. The associated change in terms of sterics—and thus fluctuations—is difficult to predict. The following simulation attempts to measure such a change.

TABLE I. Amino acid sequences of alamethicin, several WALP peptides, and magainin-H2. Underlined amino acids are non-natural: A and F refer to α-aminoisobutyric acid and C-terminal phenylalaninol, respectively[56,57] (Fig. S5 in the supplementary material[44]). All sequences form transmembrane helices.[21,56,58] The N- and C-terminus of each peptide are capped by an acetyl and *n*-methyl amide groups, respectively (except alamethicin, for which the C-terminus is embedded in the last phenylalaninol).

Name	Sequence
Alamethicin	APAAA AQAVA GLAPV AAEQF
WALP23	GWWLA LALAL ALALA LALAL WWA
WALP27	GWWLA LALAL ALALA LALAL ALALW WA
Magainin-H2	IIKKF LHSIW KFGKA FVGEI MNI

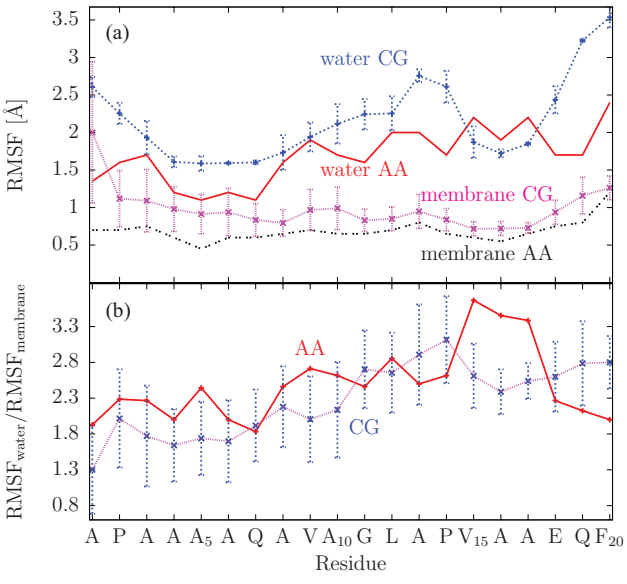

FIG. 4. (a) Root-mean-square fluctuations of alamethicin in water and in a POPC bilayer (denoted "membrane"). "AA" and "CG" correspond to atomistic and coarse-grained simulations, respectively. Atomistic data reproduced from Tieleman *et al.*[56] (b) Water/membrane RMSF ratios for atomistic ("AA") and coarse-grained ("CG").

We repeat the simulations of Tieleman *et al.*[56] using the present CG model. Non-natural amino acids were systematically replaced by natural ones: α-aminoisobutyric acid (Aib) and C-terminal phenylalaninol (Phol) into Ala and Phe, respectively (see the supplementary material[44]). Note that all simulations involving peptides (rather than single residues) required a time step $\delta t = 0.01\,\tau$.[29]

CG simulations in implicit water ($10^7\,\tau$) quickly stabilized an α-helix. The results are shown in Fig. 4(a) ("water CG"). We note that the corresponding atomistic and CG simulations in water agree reasonably well, considering that no temperature calibration was applied. The enhanced fluctuations in water at the ends of the chain (i.e., residues 1–4 and 17–20) likely result from the lack of hydrogen-bond sites at the very first amide and very last carbonyl groups, due to the lack of capping residues in the water simulation. Due to the geometry of a backbone hydrogen bond, it is likely to affect four contiguous residues. The peak around residues 12–15 illustrates the added flexibility due to both Gly_{11} and Pro_{14}. The role of proline in the conformation of the helix is explained in detail in Tieleman *et al.*[56]

CG simulations of Alm in a 72-POPC lipid membrane were run for a total simulation time $t = 500\,000\,\tau$. The initial conformation was generated by inserting a helical conformation of Alm in an equilibrated lipid bilayer. Alm did not show any significant change of secondary structure over the entire simulation (recall that it is *not* biased to stay in an α-helical conformation). As in the water case, the RMSF curves of the atomistic and CG simulations of Alm in the membrane show good agreement, except for a slight upward shift (see Fig. 4(a)). We remind the reader that no free parameter was tuned to reproduce the atomistic curves with our CG model.

Both all-atom and CG simulations show that the RMSF of the membrane-inserted helix is systematically smaller than that of the water-solubilized helix. While the RMSF of the CG simulations in both cases are slightly larger, Fig. 4(b) shows that the *ratio* $\mathrm{RMSF}_{\mathrm{water}}/\mathrm{RMSF}_{\mathrm{membrane}}$ for both resolutions is essentially the same. Since the hydrogen bond strength has not been adjusted in the CG simulations, we suspect the aligning propensity of orientationally ordered lipid tails to be responsible for the reproduction of this RMSF ratio. To avoid over-stabilization, we therefore argue that the hydrogen bond strength of the peptide model should in fact *not* be increased inside the bilayer environment. This is in line with experimental measurements suggesting that H-bond strength inside a bilayer does not in fact increase as much as one would think.[62–64]

C. Tilt and hydrophobic mismatch

The tilt angle of model transmembrane proteins—an indicator of their orientation relative to the membrane normal—has been the subject of detailed studies[65–67] in order to both better understand hydrophobic mismatch between lipids and proteins[68] as well as to test the modeling of the N- and C-termini.

CG simulations of the WALP23 and WALP27 peptides (sequences shown in Table I) were run in a 72-POPC lipid bilayer. Constant temperature ($T = T_{\mathrm{body}}$) and tension ($\Sigma = 0$) simulations were performed as above for $2 - 3 \times 10^6\,\tau$. The tilt angle ϑ was measured from the orientation of the helical principal axis (as calculated from the gyration tensor) and the unit vector along the membrane normal.

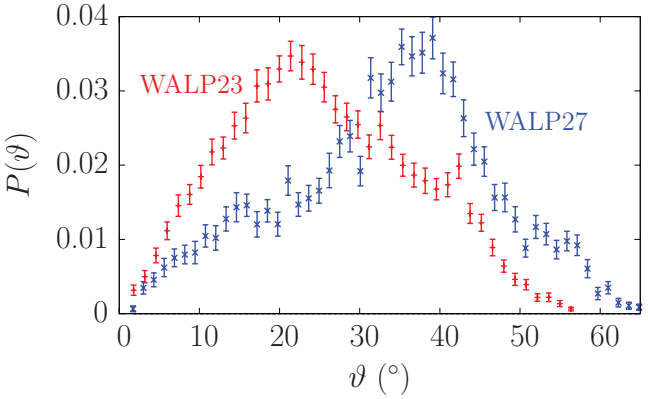

FIG. 5. Probability distribution function $P(\vartheta)$ for WALP23 (red) and WALP27 (blue).

The tilt-angle distribution $P_\vartheta(\vartheta)$ is shown in Fig. 5. The average angle $\langle\vartheta\rangle_\vartheta$ is 22° (29°) for WALP23 (WALP27). This shift for the average angle between WALP23 and WALP27 illustrates the impact of hydrophobic mismatch on the orientations of the two peptides: WALP27 being longer (i.e., having a larger positive hydrophobic mismatch), can tilt more to optimize hydrophobic matching.[68] In the supplementary material,[44] we argue from the data that this effect does *not* in fact result from a preferred tilt angle, but from a wider range of angles accessible within $k_{\mathrm{B}}T$. Our findings are in good agreement with the atomistic (umbrella sampling) simulations of Kim and Im,[58] who measured thermally accessible tilt angles in the range $7^\circ - 26^\circ$ and $14^\circ - 46^\circ$ for WALP23 and WALP27, respectively, in POPC. While using different lipids, several independent experimental and simulation studies point to an average tilt angle of $\approx 15^\circ - 25^\circ$ for WALP23 in DOPC.[69,70] Similar results were obtained in DMPC.[58,71]

D. Helix–helix interactions

The aggregation of proteins in, or close to, the lipid bilayer have important biological consequences for the membrane, for example, by triggering vesicle budding[4,72] or pore formation.[73] These phenomena depend not only on protein–lipid interactions, but also on protein–protein interactions *in the membrane environment*. The self-association of WALP peptides in model membranes—studied both experimentally and computationally[21,74]—provides an appropriate benchmark to test the cross-parametrization by studying the distance and (relative) crossing angle of WALP dimers.

We simulated WALP23 dimers in a 72-POPC lipid bilayer at constant temperature ($T = T_{\mathrm{body}}$) and vanishing tension ($\Sigma = 0$). All simulation conditions were the same as above. Two independent simulations, totaling $10^6\,\tau$, were run with helical peptides initially placed in *parallel*, integral to the bilayer, and at a 13 Å distance from one another. The helix–helix distance d was measured from the centers of mass

of the two peptides, while the crossing angle α was defined from the angle between the two helical principal axes. Fig. S8 (supplementary material[44]) shows the free energies as a function of the helix–helix distance (a) and cosine of the crossing angle (b) between the two peptides. These results compare reasonable well with atomistic simulations of WALP dimers in DOPC which report an average distance of 11 Å and angle $15^\circ - 25^\circ$[74] for two parallel dimers. We find from our simulations $\langle d \rangle = 9.5$ Å and $\langle \vartheta \rangle_\vartheta = 12.0^\circ$.

V. COOPERATIVE FEATURES OF PEPTIDE–LIPID INTERACTIONS

A. Pore formation triggered by cooperative antimicrobial peptide insertion

To illustrate the cooperative behavior of many peptides close to a lipid bilayer, we studied 8 magainin MG-H2 proteins (sequence displayed in Table I) interacting with a 288-lipid POPC membrane. We performed replica-exchange MD (REMD) simulations using 8 replicas with $T/T_{\text{body}} \in [1.0, 1.29]$ to circumvent kinetic trapping. Each was run for $300\,000\,\tau$. The initial conformation was generated by combining a pre-equilibrated stress-free lipid bilayer ($\Sigma = 0$; $T = T_{\text{body}}$) and 8 Magainin MG-H2 peptides. The peptides were initially placed in solution at a short distance of 12 Å from the boundary of the lipid membrane, half of them above and the other half underneath the membrane (see Fig. 6). The peptide–lipid ratio (36 lipids per peptide) is comparable to existing all-atom[75] and CG[21] simulations, as well as experimental condition.[76] The initial conformation of each magainin peptide was chosen as a folded α-helix rather than an unfolded coil, in line with previous all-atom[75] and CG simulations.[21]

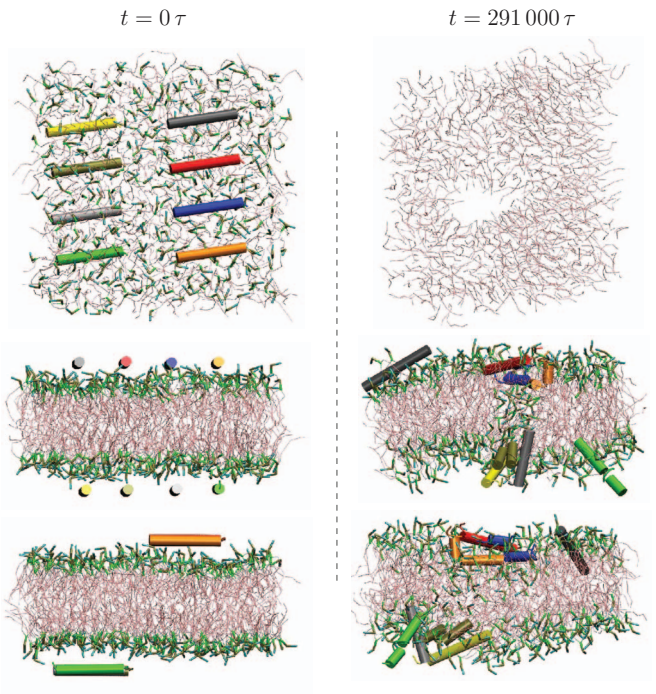

FIG. 6. CG REMD simulation replica of the cooperative insertion of magainin MG-H2 peptides in a 288-POPC-lipid membrane bilayer at $T = 1.025\,T_{\text{body}}$. Left: top and two side views of the initial configuration ($t = 0$). Right: top (along the pore direction; only the lipid tails are shown for clarity) and two side views of the final configuration ($t = 291\,000\,\tau$). Lipid beads are colored according to their type: CG (cyan), PH (tan), GL (green), ES (silver), alkyl tails (pink). Each protein is colored differently; cylindrical motifs indicate a helical conformation.

Our simulations show that although the magainin peptides rapidly adsorb onto the membrane, a small amount of unfolding takes place while the peptides are still in solution, which leads to partially unfolded peptides after membrane binding. We reiterate that the partially unfolded conformation is consistent with state of the art all-atom simulations of magainin peptides interacting with a membrane[75] and CD experiments.[76,77]

The partially unfolded peptides tend to cluster together while interacting with the membrane and their aggregation promotes the formation of a pore, as can be seen on the right-hand side of Fig. 6 ($T = 1.025\,T_{\text{body}}$). Both the peptides' conformations and orientations vary widely during the CG simulations. Our simulations show a disordered toroidal pore, a structure reminiscent of what is observed in all-atom simulations,[75] where the pore is not cylindrical and the MG-H2 peptides prefer to stay close to the edge of the pore rather than perpendicular to the plane of the membrane. As seen in Fig. 6, the formation of a pore is clearly a cooperative process facilitated by peptide aggregation. The two groups of four MG-H2 on both sides of the membrane were initially not placed exactly opposite of each other but with a small lateral offset (see the left-hand side of Fig. 6). Therefore, the nascent toroidal pore generated by the peptides follows the tilted direction formed by these two groups of adsorbed peptides, rather than being strictly perpendicular to the membrane plane (see the right-hand side of Fig. 6). Interestingly, the formation of a pore does not require the penetration of any MG-H2 *through* the bilayer, in agreement with all-atom simulations.[75]

Our short simulations are of course insufficient to elucidate the pore formation process of magainin. But this was not our aim. We rather wanted to show that our cross-parametrized model can access the length and time scales relevant for the large-scale structural rearrangements involved in this process, and that the results which our simple test simulations show are compatible with what is currently known. This justifies the use of such models for more extensive studies, while of course always bearing in mind the limitations inherent in such a CG approach.

B. Insertion and folding of WALP23

Proteins are structurally stabilized not only by the nature of their intramolecular interactions but also by the surrounding environment, such as water or a lipid bilayer.[61] Since our peptide model does not explicitly constrain secondary structure, we explore its ability to fold a simple peptide, WALP23, in both water and the membrane.

We first performed REMD simulations of WALP23 in pure water, which resulted in very low helicity—less than 5%—throughout the sequence at temperatures $T/T_{\text{body}} \in [1.0, 1.5]$ (see Sec. III A in the supplementary material[44]). This

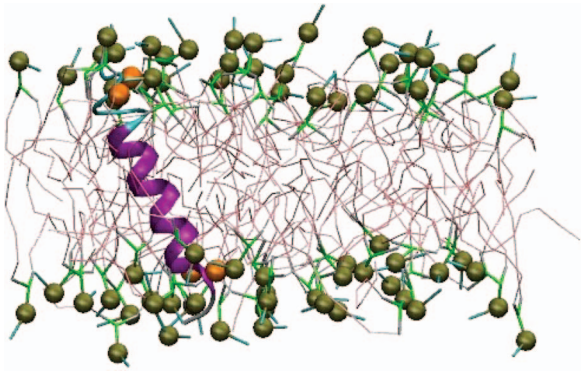

FIG. 7. Folded WALP23 conformation, sampled at $T = 1.15\,T_{\text{body}}$. Lipid beads are colored according to their type (see Fig. 6). The orange beads represent the tryptophan residues at each end of WALP23. The rest of the protein is represented according to its secondary structure: turn (green), α helix (purple), π helix (red), coil (white).

absence of α-helix formation is in agreement with experimental and all-atom simulation results.[78–80]

We subsequently ran REMD simulations of the same peptide, initially unstructured and placed close to—but out of—a 72 POPC lipid membrane. The temperature range, $T/T_{\text{body}} \in [1.0, 1.2]$, allowed our lipid bilayer to remain in a stable fluid phase even at the highest temperature,[35] thus eliminating the need to artificially stabilize the membrane.[78–80] For more details on our REMD setup, see Sec. III B in the supplementary material.[44]

While our REMD protocol provided reasonable energy overlaps and acceptance ratios, it failed to successfully couple the entire range of temperature, instead coupling each replica only with its nearest neighbor(s) (see the supplementary material[44])—pointing to the inadequacy of a standard REMD protocol for this system. By the end of the simulation, the two highest temperature replicas (i.e., $T/T_{\text{body}} = 1.15$ and 1.2) had inserted and folded WALP23 as a transmembrane α-helix (see Fig. 7 for a folded conformation and Fig. S11 for representative snapshots of the folding event). The three colder replicas, on the other hand, failed to show a folded helix, posing the question whether this is for thermodynamic or kinetic reasons. The lack of coupling between the set of replicas prevents us from making a definite statement, though we provide in the following a number of arguments in favor of the latter reason. Given enough simulation time a single folded helix should form since this process is largely enthalpy-driven.[81] Indeed, the helicity decreases rather than increases with temperature, as also shown in our simulations of WALP23 folding in solvent (see Fig. S9 in the supplementary material[44]). In addition, we observed that α-helical folds formed in our CG peptide model have a tendency to remain stable over a wide temperature interval below the unfolding temperature.[31] Finally, we ran another REMD simulation starting from the folded (i.e., integral and helical) conformation, as observed at the highest temperatures. After a simulation time of $t = 1.2 \times 10^6\,\tau$, none of the replicas unfolded (Fig. S12 in the supplementary material[44]). Peptides in all replicas remained integral to the membrane and kept a fractional helicity of more than 70%. If in the absence of kinetic issues the folded state at $T = T_{\text{body}}$ were not the equilibrium state, then these simulations should unfold the peptide. This is not observed, and taken together with the arguments from above this provides further support that the integral helix is indeed the thermodynamically stable state near physiological temperatures in this model.

If one agrees with this interpretation, then the failure to achieve the integral helical conformation is a sobering reminder how very slow such folding processes in a crowded and ordered membrane environment can be, even for as simple a peptide as WALP23. The smallness of the peptide and the ease with which systems of this size reach thermal equilibrium in bulk water hides the fact that its free energy landscape within a lipid bilayer might be much more complex. Our failure to fold WALP23 does of course not actually prove this, and more replicas and perhaps finer refined algorithmic techniques (e.g., Hamiltonian Replica exchange MD[82]) might be necessary to investigate this problem in more detail. This, however, goes beyond what we wish to show in the present paper.

VI. CONCLUSIONS

We have seen that a fairly simple cross-parametrization of an implicit solvent CG peptide model with an implicit solvent CG membrane model goes a surprisingly long way in reproducing a number of interesting peptide-membrane phenomena: structural properties commonly tested in CG force-field papers[21,24] (e.g., transmembrane-helix fluctuations, tilt, and helix-helix distance and angle), aggregation, and structure formation. However, the underlying biophysical problems are highly complex, and one cannot expect with such a simple approach to have arrived at a model that faithfully captures them all. Let us therefore conclude with a number of limitations that, in our view, constitute important caveats to bear in mind when applying models such as the one we discussed here:

- Reproducing atomistic PMFs of the insertion of single amino-acid side-chains in a DOPC bilayer[41] provides a parametrization scheme pertaining to the energetics of peptide–lipid interactions. Evidently, the resulting parametrization is not unique: a number of parameters can reproduce the observables of interest with the same accuracy. Resolving such an under-determinacy situation would lead to slightly different models. In addition, one should bear in mind that a variety of sources of errors are intrinsically present: atomistic force fields; sampling of the atomistic simulations;[83] CG parametrization protocol; CG free-energy calculations.
- The cross-parametrization is inappropriate for phenomena that heavily rely on electrostatics (e.g., electroporation[84]). The two original peptide[29] and lipid[35] models do not incorporate explicit $1/r$ electrostatics. While frequently essential, electrostatics is not only computationally expensive, but—and this is the much more severe problem—very difficult to systematically coarse-grain. We point out that (i) simply putting charges on the beads and working out $1/r$

interactions is not the right thing to do in the presence of dielectric discontinuities, (ii) the CG lipid model used here targets neutral lipids—the contribution from partial charges being, to a large extent, absorbed in the CG potentials, and (iii) the interaction between charged amino acids is somewhat reproduced in the Miyazawa-Jernigan matrix[85] used for the side-chain–side-chain interactions.

- The lipid–peptide cross-parametrization was only calibrated at body temperature in the (membrane) fluid phase. Although the lipid model was systematically derived based on structure[35] and has demonstrated lipid and phase transferability in terms of structure and area per lipid,[36] temperature and phase transferability has not been investigated for the cross-interactions.
- While the folding of WALP23 was only observed above body temperature (i.e., $T = 1.15 - 1.20\ k_B T_{body}$), the fact that it was indeed possible to fold a helix from scratch is nevertheless encouraging. Bearing in mind the prominent strength of hydrogen bonds in the bilayer environment, the α-helix is the secondary structure of choice in most proteins. Even though helix formation is overall strongly favored, the model seems to discriminate the amount of helicity of WALP depending on its environment: none in the water, more helical in the membrane. The difficulties we have encountered in folding the helix in the membrane underline the challenges associated with peptide structure formation in a lipid environment. We suspect that the underlying problem in achieving the helical state at body temperature are kinetic in nature, and this in itself should be a warning: Simply because one has switched to a lower resolution CG model and employed parallel tempering simulations is not by itself a guarantee that one will now readily achieve thermal equilibrium.
- As of yet, the model is out of reach of tertiary structure format folding in the membrane, which would require a more refined parametrization and further testing, though simple structural tests of helix-helix interactions (i.e., distance and angle) look promising. Specifically, the crude description of electrostatics and the approximate parametrization of the N- and C-termini are shortcomings that would need to be more thoroughly addressed.

ACKNOWLEDGMENTS

We thank C. Peter, C. Junghans, P. Fuchs, and L. Monticelli for many insightful discussions.

This work was partially supported by the National Institutes of Health (NIH) (Grant No. P01AG032131). T.B. received an Astrid and Bruce McWilliams Fellowship. The computations were supported in part by the National Science Foundation (NSF) through TeraGrid resources provided by the Pittsburgh Supercomputing Center (PSC) and The Texas Advanced Computing Center (TACC).

[1] J. Lyklema, *Fundamentals of Interface and Colloid Science: Soft Colloids* (Academic Press, 2005), Vol. 5.
[2] D. M. Engelman, "Membranes are more mosaic than fluid," Nature (London) **438**, 578–580 (2005).
[3] M. Matsumoto, S. Saito, and I. Ohmine, "Molecular dynamics simulation of the ice nucleation and growth process leading to water freezing," Nature (London) **416**(6879), 409–413 (2002).
[4] P. D. Blood and G. A. Voth, "Direct observation of Bin/amphiphysin/Rvs (BAR) domain-induced membrane curvature by means of molecular dynamics simulations," Proc. Natl. Acad. Sci. U.S.A. **103**(41), 15068–15072 (2006).
[5] J. Gumbart, L. G. Trabuco, E. Schreiner, E. Villa, and K. Schulten, "Regulation of the protein-conducting channel by a bound ribosome," Structure **17**(11), 1453–1464 (2009).
[6] D. S. D. Larsson, L. Liljas, and D. van der Spoel, "Virus capsid dissolution studied by microsecond molecular dynamics simulations," PLoS Comput. Biol. **8**(5), e1002502 (2012).
[7] G. Zhao, J. R. Perilla, E. L. Yufenyuy, X. Meng, B. Chen, J. Ning, J. Ahn, A. M. Gronenborn, K. Schulten, C. Aiken, and P. Zhang, "Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics," Nature (London) **497**, 643–646 (2013).
[8] D. W. Borhani and D. E. Shaw, "The future of molecular dynamics simulations in drug discovery," J. Comput.-Aided Mol. Des. **26**(1), 15–26 (2012).
[9] G. A. Voth, *Coarse-Graining of Condensed Phase and Biomolecular Systems*, 1st ed. (CRC Press, 2008).
[10] T. Murtola *et al.*, "Multiscale modeling of emergent materials: Biological and soft matter," in *Special issue on: Coarse-grained modeling of soft condensed matter*, Phys. Chem. Chem. Phys. **11**, 1869–1892 (2009).
[11] J. J. de Pablo, "Coarse-grained simulations of macromolecules: From DNA to nanocomposites," Annu. Rev. Phys. Chem. **62**, 555–574 (2011).
[12] W. G. Noid, "Perspective: Coarse-grained models for biomolecular systems," J. Chem. Phys. **139**(9), 090901 (2013).
[13] D. Yin and A. D. MacKerell, "Combined *ab initio*/empirical approach for optimization of Lennard–Jones parameters," J. Comput. Chem. **19**(3), 334–348 (1998).
[14] W. G. Noid, "Systematic methods for structurally consistent coarse-grained models," *Biomolecular Simulations* (Springer, 2013), pp. 487–531.
[15] M. Jochum, D. Andrienko, K. Kremer, and C. Peter, "Structure-based coarse-graining in liquid slabs," J. Chem. Phys. **137**, 064102 (2012).
[16] M. Venturoli, B. Smit, and M. M. Sperotto, "Simulation studies of protein-induced bilayer deformations, and lipid-induced protein tilting, on a mesoscopic model for lipid bilayers with embedded proteins," Biophys. J. **88**(3), 1778–1798 (2005).
[17] P. J. Bond and M. S. P. Sansom, "Insertion and assembly of membrane proteins via simulation," J. Am. Chem. Soc. **128**(8), 2697–2704 (2006).
[18] Q. Shi, S. Izvekov, and G. A. Voth, "Mixed atomistic and coarse-grained molecular dynamics: Simulation of a membrane-bound ion channel," J. Phys. Chem. B **110**(31), 15045–15048 (2006).
[19] A. Y. Shih, A. Arkhipov, P. L. Freddolino, and K. Schulten, "Coarse grained protein-lipid model with application to lipoprotein particles," J. Phys. Chem. B **110**(8), 3674–3684 (2006).
[20] G. Illya and M. Deserno, "Coarse-grained simulation studies of peptide-induced pore formation," Biophys. J. **95**(9), 4163–4173 (2008).
[21] L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, and S. J. Marrink, "The Martini coarse-grained force field: Extension to proteins," J. Chem. Theory Comput. **4**(5), 819–834 (2008).
[22] B. West, F. L. H. Brown, and F. Schmid, "Membrane-protein interactions in a generic coarse-grained model for lipid bilayers," Biophys. J. **96**(1), 101–115 (2009).
[23] Z. Wu, Q. Cui, and A. Yethiraj, "A new coarse-grained force field for membrane-peptide simulations," J. Chem. Theory Comput. **7**(11), 3793–3802 (2011).
[24] C.-K. Wan, W. Han, and Y.-D. Wu, "Parameterization of pace force field for membrane environment and simulation of helical peptides and helix–helix association," J. Chem. Theory Comput. **8**(1), 300–313 (2012).
[25] A. Dryga, S. Chakrabarty, S. Vicatos, and A. Warshel, "Coarse grained model for exploring voltage dependent ion channels," BBA - Biomembranes **1818**(2), 303–317 (2012).
[26] S. J. Marrink and D. P. Tieleman, "Perspective on the martini model," Chem. Soc. Rev. **42**, 6801–6822 (2013).
[27] J. W. Mullinax and W. G. Noid, "Extended ensemble approach for deriving transferable coarse-grained potentials," J. Chem. Phys. **131**, 104110 (2009).
[28] R. C. DeMille, T. E. Cheatham III, and V. Molinero, "A coarse-grained model of DNA with explicit solvation by water and ions," J. Phys. Chem. B **115**(1), 132–142 (2011).

[29]T. Bereau and M. Deserno, "Generic coarse-grained model for protein folding and aggregation," J. Chem. Phys. **130**(23), 235106–235120 (2009).
[30]T. Bereau, M. Bachmann, and M. Deserno, "Interplay between secondary and tertiary structure formation in protein folding cooperativity," J. Am. Chem. Soc. **132**(38), 13129–13131 (2010).
[31]T. Bereau, M. Deserno, and M. Bachmann, "Structural basis of folding cooperativity in model proteins: Insights from a microcanonical perspective," Biophys. J. **100**(11), 2764–2772 (2011).
[32]T. Bereau, C. Globisch, M. Deserno, and C. Peter, "Coarse-grained and atomistic simulations of the salt-stable cowpea chlorotic mottle virus (SS-CCMV) subunit 26–49: β-barrel stability of the hexamer and pentamer geometries," J. Chem. Theory Comput. **8**, 3750–3758 (2012).
[33]X. Periole, M. Cavalli, S. J. Marrink, and M. Ceruso, "Combining an elastic network with a coarse-grained molecular force field: Structure, dynamics and intermolecular recognition," J. Chem. Theory Comput. **5**, 2531–2543 (2009).
[34]C. Globisch, V. Krishnamani, M. Deserno, and C. Peter, "Optimization of an elastic network augmented coarse grained model to study CCMV capsid deformation," PloS ONE **8**(4), e60582 (2013).
[35]Z.-J. Wang and M. Deserno, "A systematically coarse-grained solvent-free model for quantitative phospholipid bilayer simulations," J. Phys. Chem. B **114**(34), 11207–11220 (2010).
[36]Z.-J. Wang and M. Deserno, "Systematic implicit solvent coarse-graining of bilayer membranes: Lipid and phase transferability of the force field," New J. Phys. **12**, 095004 (2010).
[37]A. K. Soper, "Empirical potential Monte Carlo simulation of fluid structure," Chem. Phys. **202**(2–3), 295–306 (1996).
[38]D. Reith, M. Pütz, and F. Müller-Plathe, "Deriving effective mesoscale potentials from atomistic simulations," J. Comput. Chem. **24**(13), 1624–1636 (2003).
[39]J. L. Fauchere and V. Pliska, "Hydrophobic parameters pi of amino-acid side-chains from the partitioning of *n*-acetyl-amino-acid amides," Eur. J. Med. Chem. **18**(4), 369–375 (1983).
[40]W. C. Wimley and S. H. White, "Experimentally determined hydrophobicity scale for proteins at membrane interfaces," Nat. Struct. Biol. **3**, 842–848 (1996).
[41]J. L. MacCallum, W. F. D. Bennett, and D. P. Tieleman, "Distribution of amino acids in a lipid bilayer from computer simulations," Biophys. J. **94**(9), 3393–3404 (2008).
[42]W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, "Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids," J. Am. Chem. Soc. **118**(45), 11225–11236 (1996).
[43]G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, and W. L. Jorgensen, "Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides," J. Phys. Chem. B **105**(28), 6474–6487 (2001).
[44]See supplementary material at http://dx.doi.org/10.1063/1.4867465 for more details on the cross-parametrization; modeling of the termini; validation from transmembrane helices (fluctuations, tilt, and helix–helix interactions); WALP23 simulations in water and the membrane.
[45]R. L. Henderson, "A uniqueness theorem for fluid pair correlation functions," Phys. Lett. A **49**(3), 197–198 (1974).
[46]W. Humphrey, A. Dalke, and K. Schulten, "VMD: Visual molecular dynamics," J. Mol. Graph. **14**(1), 33–38 (1996).
[47]G. M. Torrie and J. P. Valleau, "Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling," J. Comput. Phys. **23**(2), 187–199 (1977).
[48]Constraining two particles at different *heights* might avoid significant structural effects on the bilayer.
[49]A. M. Ferrenberg and R. H. Swendsen, "Optimized Monte Carlo data analysis," Phys. Rev. Lett. **63**(12), 1195–1198 (1989).
[50]S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, "The weighted histogram analysis method for free-energy calculations on biomolecules: I. The method," J. Comput. Chem. **13**(8), 1011–1021 (1992).
[51]T. Bereau and R. H. Swendsen, "Optimized convergence for multiple histogram analysis," J. Comput. Phys. **228**(17), 6119–6129 (2009).
[52]M. R. Chernick, *Bootstrap Methods: A Guide for Practitioners and Researchers*, 2nd ed. (Wiley-Interscience, 2008).
[53]L. Monticelli, K. M. Robertson, J. L. MacCallum, and D. P. Tieleman, "Computer simulation of the KvAP voltage-gated potassium channel: Steered molecular dynamics of the voltage sensor," FEBS Lett. **564**(3), 325–332 (2004).
[54]J. A. Freites, D. J. Tobias, G. von Heijne, and S. H. White, "Interface connections of a transmembrane voltage sensor," Proc. Natl. Acad. Sci. U.S.A. **102**(42), 15059–15064 (2005).
[55]I. Vorobyov, L. Li, and T. W. Allen, "Assessing atomistic and coarse-grained force fields for protein-lipid interactions: The formidable challenge of an ionizable side chain in a membrane," J. Phys. Chem. B **112**(32), 9588–9602 (2008).
[56]D. P. Tieleman, M. S. P. Sansom, and H. J. C. Berendsen, "Alamethicin helices in a bilayer and in solution: Molecular dynamics simulations," Biophys. J. **76**(1), 40–49 (1999).
[57]R. Improta, V. Barone, K. N. Kudin, and G. E. Scuseria, "Structure and conformational behavior of biopolymers by density functional calculations employing periodic boundary conditions. I. The case of polyglycine, polyalanine, and poly-α-aminoisobutyric acid *in vacuo*," J. Am. Chem. Soc. **123**, 3311–3322 (2001).
[58]T. Kim and W. Im, "Revisiting hydrophobic mismatch with free energy simulation studies of transmembrane helix tilt and rotation," Biophys. J. **99**, 175–183 (2010).
[59]K. Kuczera, "Dynamics and thermodynamics of globins," in *Recent Developments in Theoretical Studies of Proteins*, Advanced Series in Physical Chemistry Vol. 7, edited by R. Elber (World Scientific, 1996), pp. 1–64.
[60]T. Bereau, "Unconstrained structure formation in coarse-grained protein simulations," Ph.D. thesis, Carnegie Mellon University, Pittsburgh, PA, USA, 2011.
[61]S. H. White and W. C. Wimley, "Membrane protein folding and stability: Physical principles," Annu. Rev. Biophys. Biomol. Struct. **28**, 319–365 (1999).
[62]G. Grigoryan and W. F. DeGrado, "Modest membrane hydrogen bonds deliver rich results," Nat. Chem. Biol. **4**, 393–394 (2008).
[63]N. H. Joh, A. Min, S. Faham, J. P. Whitelegge, D. Yang, V. L. Woods, and J. U. Bowie, "Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins," Nature (London) **453**, 1266–1270 (2008).
[64]J. U. Bowie, "Membrane protein folding: How important are hydrogen bonds?," Curr. Opin. Struct. Biol. **21**(1), 42–49 (2011).
[65]M. R. R. de Planque and J. Antoinette Killian, "Protein–lipid interactions studied with designed transmembrane peptides: Role of hydrophobic matching and interfacial anchoring (review)," Mol. Membr. Biol. **20**(4), 271–284 (2003).
[66]S. H. Park and S. J. Opella, "Tilt angle of a trans-membrane helix is determined by hydrophobic mismatch," J. Mol. Biol. **350**(2), 310–318 (2005).
[67]M. B. Ulmschneider, M. S. P. Sansom, and A. Di Nola, "Evaluating tilt angles of membrane-associated helices: Comparison of computational and NMR techniques," Biophys. J. **90**(5), 1650–1660 (2006).
[68]O. G. Mouritsen and M. Bloom, "Mattress model of lipid-protein interactions in membranes," Biophys. J. **46**, 141–153 (1984).
[69]L. Monticelli, D. P. Tieleman, and P. F. J. Fuchs, "Interpretation of ^{2}H-NMR experiments on the orientation of the transmembrane helix WALP23 by computer simulations," Biophys. J. **99**, 1455–1464 (2010).
[70]A. Holt, R. B. M. Koehorst, T. Rutters-Meijneke, M. H. Gelb, D. T. S. Rijkers, M. A. Hemminga, and J. A. Killian, "Tilt and rotation angles of a transmembrane model peptide as studied by fluorescence spectroscopy," Biophys. J. **97**, 2258–2266 (2009).
[71]S. Özdirekcan, C. Etchebest, J. A. Killian, and P. F. J. Fuchs, "On the orientation of a designed transmembrane peptide: Toward the right tilt angle?," J. Am. Chem. Soc. **129**, 15174–15181 (2007).
[72]B. J. Reynwar, G. Illya, V. A. Harmandaris, M. M. Müller, K. Kremer, and M. Deserno, "Aggregation and vesiculation of membrane proteins by curvature-mediated interactions," Nature (London) **447**, 461–464 (2007).
[73]Z. Oren and Y. Shai, "Mode of action of linear amphipathic α-helical antimicrobial peptides," Peptide Sci. **47**, 451–463 (1998).
[74]E. Sparr, W. L. Ash, P. V. Nazarov, D. T. Rijkers, M. A. Hemminga, D. P. Tieleman, and J. A. Killian, "Self-association of transmembrane alpha-helices in model membranes: Importance of helix orientation and role of hydrophobic mismatch," J. Biol. Chem. **280**, 39324–39331 (2005).
[75]H. Leontiadou, A. E. Mark, and S. J. Marrink, "Antimicrobial peptides in action," J. Am. Chem. Soc. **128**(37), 12156–12161 (2006).
[76]T. Tachi, R. F. Epand, R. M. Epand, and K. Matsuzaki, "Position-dependent hydrophobicity of the antimicrobial magainin peptide affects the mode of peptide-lipid interactions and selective toxicity," Biochemistry **41**(34), 10723–10731 (2002).

[77]B. Bechinger, "Structure and function of membrane-lytic peptides," Crit. Rev. Plant Sci. **23**(3), 271–292 (2004).
[78]M. B. Ulmschneider, J. P. F. Doux, J. Antoinette Killian, J. C. Smith, and J. P. Ulmschneider, "Mechanism and kinetics of peptide partitioning into membranes from all-atom simulations of thermostable peptides," J. Am. Chem. Soc. **132**(10), 3452–3460 (2010).
[79]H. Nymeyer, T. B. Woolf, and A. E. Garcia, "Folding is not required for bilayer insertion: Replica exchange simulations of an α-helical peptide with an explicit lipid bilayer," Proteins: Struct. Funct. Bioinf. **59**, 783–790 (2005).
[80]M. B. Ulmschneider and J. P. Ulmschneider, "Folding peptides into lipid bilayer membranes," J. Chem. Theory Comput. **4**(11), 1807–1809 (2008).
[81]T. Ooi and M. Oobatake, "Prediction of the thermodynamics of protein unfolding: The helix-coil transition of poly(l-alanine)," Proc. Natl. Acad. Sci. U.S.A. **88**(7), 2859–2863 (1991).
[82]A. Bunker and B. Dünweg, "Parallel excluded volume tempering for polymer melts," Phys. Rev. E **63**(1), 016701 (2000).
[83]C. Neale, W. F. D. Bennett, D. P. Tieleman, and R. Pomès, "Statistical convergence of equilibrium properties in simulations of molecular solutes embedded in lipid bilayers," J. Chem. Theory Comput. **7**(12), 4175–4188 (2011).
[84]D. P. Tieleman, "The molecular basis of electroporation," BMC Biochem. **5**, 10 (2004).
[85]S. Miyazawa and R. L. Jernigan, "Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading," J. Mol. Biol. **256**(3), 623–644 (1996).